DOCUMENT RESUME

ED 437 834 FL 026 109

AUTHOR Fortkamp, Maike B. Mota
TITLE Working Memory Capacity and L2 Speech Production.
PUB DATE 1999-00-00
NOTE 36p.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Articulation (Speech); Cognitive Measurement; *Cognitive Processes; English (Second Language); *Language Fluency; Language Tests; *Memory; *Oral language; Portuguese; *Psycholinguistics; *Second Language Learning; Second Languages; Testing

ABSTRACT

This study examines whether working memory capacity, a construct of current information processing theory, correlates with fluent foreign language (L2) speech production. It is based on M. Daneman's study (1991), who found significant correlation between individuals' working memory capacity and the fluency with which they can speak their first language (L1). Adapting Daneman's methodology, a set of seven experiments was applied to 16 advanced speakers of English as a foreign language. Working memory was assessed by means of the Speaking Span Test and the Reading Span Test, both in Portuguese and English. L2 fluency was assessed by means of the Speech Generation Task, which was aimed at assessing fluency at the discourse level. Working memory capacity both in Portuguese and English correlated significantly only with the reading-related task; the Oral Reading Task, aimed at assessing fluency at the articulatory level. The results of this test support the task-specific view of working memory capacity, which posits that this capacity is functional, varying according to the individual's efficiency in the processes specific to the cognitive task with which it is correlated. Various figures, graphs, and charts appear throughout the body of the work. (Contains 65 references.) (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 437 834

Working memory capacity and L2 speech production

Mailce B. Mota Fortkamp
Universidade Federal de Santa Catarina

Abstract

This study examines whether working memory capacity, a construct of current information processing theory, correlates with fluent foreign language (L2) speech production. It is based on Daneman (1991), who found a significant correlation between individuals' working memory capacity and the fluency with which they can speak in their first language (L1). Adapting Daneman's (1991) methodology, a set of seven experiments was applied to 16 advanced speakers of English as a foreign language. Working memory was assessed by means of the Speaking Span Test (Daneman & Green, 1986; Daneman, 1991) and the Reading Span Test (Daneman & Carpenter, 1980 and 1983), both in Portuguese and English. L2 fluency was assessed by means of the Speech Generation Task, the Oral Reading Task, and the Oral Slip Task (Motley and Baars, 1976). Working memory capacity, as measured by the Speaking Span Test in English, correlated significantly only with the Speech Generation Task, which was aimed at assessing fluency at the discourse level. Working memory capacity, as measured by the Reading Span Test, both in Portuguese and English, correlated significantly only with the reading-related task, the Oral Reading Task, aimed at assessing fluency at the articulatory level. The results of the present study support the task-specific view of working memory capacity (see Cantor and Engle, 1993), which posits that this capacity is functional, varying according to the individual's efficiency in the processes specific to the cognitive task with which it is being correlated.

In the past few years cognitive processes involved in second/foreign language (L2) acquisition and use have gained increased importance in L2 acquisition/use research. To gain insights into the relationship between cognition and L2 acquisition/use, researchers in the L2 acquisition/use area have drawn on studies developed in the cognitive sciences, which seek to understand and explain the mental processes involved in a number of tasks such as perceiving, remembering, understanding, learning, and reasoning (Ashcraft, 1994; Stillings, Feinstein, Garfield, Rissland, Rosenbaum, Weisler, & Baker-Ward, 1987).

The integration of cognitive models within L2 studies has been limited to certain aspects of the L2, with research accumulating in the area of language comprehension--mainly reading--and only a few studies in language production. Among the latter, the focus has been mainly on the phonological aspect of the L2 and the literature mentions only a few attempts at describing, from the cognitive perspective, the process of L2 speech production (e.g., Faerch & Kasper, 1983; Dechert & Raupach, 1980; Dechert and Raupach, 1987; de Bot, 1992).

The objective of the present article is to examine the relationship between working memory capacity, a construct of current information processing theory, and fluent L2 speech production at the discourse level and the articulatory level. As Miyake and Friedman (1998) point out, the concept of working memory may be useful in explaining individual differences in the acquisition and use of an L2. By better understanding variation in L2 performance, researchers in the area may refine theories of L2 acquisition/use and optimize the outcome of L2 instruction.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Malice B. Mota Fortkamp
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

FL026109

This article is organized in 4 sections. In the first section, the concept of working memory is discussed and a review of the psychometric correlational approach –focusing mainly on the relationship between working memory capacity and L1 processing- is presented. In the second section, relevant studies on L2 speech production are reviewed. In the third section, the method, materials, and tasks used in this study are presented, followed by the analysis and discussion of the results obtained. Finally, in section 4, the limitations of the study are outlined and suggestions for further research are made.

I Working memory

Working memory is the human limited capacity cognitive system responsible for the temporary storage and processing of information retrieved from long-term memory in the performance of complex cognitive tasks (Baddeley, 1990, 1999; Daneman, 1991; Engle, 1996; Logie, 1996; Richardson, 1996).

As Baddeley (1992) suggests, research on working memory has been developed along two different but complementary approaches. The first one, the dual-task neuropsychological approach, focuses on the analysis of the structure of the three-component working memory model proposed by Baddeley and Hitch (1974) and Hitch and Baddeley (1976). This model consists of a general-purpose central control architecture –the central executive- and two slave subsystems – the phonological loop and the visuo-spatial sketchpad. The methodology of the dual-task approach consists of the application of dual tasks--for instance, remembering a list of digits while reasoning (Baddeley, 1990:68)-- and the study of neuropsychological evidence to explain, mainly, the slave subsystems.

The second approach, the psychometric correlational, is concerned with the correlations existing between working memory capacity –conceptualized as a single unitary device- and the performance of complex cognitive tasks. Within this approach, the two functions of working memory--storage and processing--compete for its capacity during the performance of complex cognitive tasks (Daneman and Carpenter, 1980, 1983). The methodology generally consists of devising laboratory tasks in which both storage and processing of information are necessary, and subsequently using the individual's results of performance on these tasks to predict his/her skills in demanding cognitive tasks, such as reading comprehension. The present study was carried out within the psychometric correlational approach and focuses on L2 speech production.

The psychometric correlational approach to working memory capacity

Under the assumption that working memory has the dual function of storing and processing information and that traditional digit or word span tasks do not reflect the processing function efficiently, Daneman and Carpenter (1980) devised a complex measure of working memory span which they termed the Reading Span Test. In their view, there is a trade-off between storage and processing in working memory, which is likely to be a source of individual differences in reading comprehension. They propose (Daneman and Carpenter, 1980, 1983), then, that the processing and storage functions of working memory compete for its limited capacity.

The Reading Span Test, as it was first devised by Daneman and Carpenter (1980), requires subjects to use both functions of working memory: the processing component requires sentence comprehension while the storage component consists of maintaining and retrieving the final word of each sentence of a presented set. A subject's reading span is the maximum number of sentence-final words recalled in the order they were presented and is taken as an index of his/her working memory capacity.

The Reading Span Test has been the basis for most of the research on individual differences in working memory and reading comprehension, and has been extensively used as a predictor of performance on various other aspects of reading: (1) the ability to detect inconsistencies in sentences with homonyms (Daneman & Carpenter, 1983); (2) the ability subjects have to make inferences of ideas not explicitly mentioned in the text (Masson & Miller, 1983); (3) the ability to make use of contextual cues to infer the meaning of new words in the text (Daneman & Green, 1986); (4) the resolution of lexical ambiguity in reading (Miyake, Just & Carpenter, 1994); and (5) the perception of text structure (Tomitch, 1995). Various researchers (e.g. Daneman & Carpenter, 1980; Masson & Miller, 1983; Turner & Engle, 1989) have also found strong correlations between the Reading Span Test and standardized measures of reading ability such as the Verbal Scholastic Aptitude Test and the Nelson-Denny reading test.[1]

Claiming that individuals differ considerably in the fluency with which they speak, Daneman (1991)--on which the present study is based--verified whether differences in working memory capacity could account for this variation in L1 speech production. Considering that speaking is a complex cognitive task which requires coordination of storage and processing of information in the various stages of the speech production process, Daneman hypothesized that individuals with larger working memory capacities would perform better on tasks measuring fluency.

---

[1] For an extensive review of the literature on individual differences in working memory capacity and reading comprehension, see Tomitch 1995.

Subjects' working memory capacity was assessed by means of the Speaking Span Test (Daneman & Green, 1986), aimed at measuring working memory capacity during speech production. The test consisted of presenting subjects with increasingly longer sets of unrelated words, which they had to read silently. At the end of a set, subjects were required to produce aloud a sentence for each individual word presented, in their original order and form of presentation. A subject's speaking span was operationalized in terms of his/her total capacity--the total number of words for which he/she was able to produce a grammatical sentence. This total capacity was expressed in two speaking span scores: speaking span strict, counting only those sentences with the exact form of the word presented, and the speaking span lenient, counting also sentences containing the word in a different form.

Subjects' oral fluency was assessed by means of the Speech Generation Task, the Oral Reading Task and the Oral Slip Task. The Speech Generation Task aimed at eliciting fluency at the discourse level and consisted of the description of a picture for 1 minute and 30 seconds. Measures of fluency in this task were number of words completely articulated--the main measure--and richness and originality of context. The Oral Reading Task and the Oral Slip Task both aimed at eliciting fluency in terms of speed and accuracy in the articulation of words. In the Oral Reading Task subjects were required to read a passage aloud and the main measure of fluency was reading time. In the Oral Slip Task, which aimed at eliciting spoonerisms, subjects were required to say cued pairs of words shown on a computer screen. The measures of fluency used were number and types of errors made. In addition, Daneman applied a Reading Span Test, which she hypothesized would correlate with the reading related task.

The study was carried out with 29 English L1 university students and results show that the Speaking Span Test correlated significantly with the Speech Generation Task, The Oral Reading Task, and the Oral Slip Task--that is, subjects with larger working memory capacity performed better on the picture description task, took less time reading the passage aloud, and were less prone to producing spoonerisms. Also, as predicted, the Reading Span Test correlated significantly only with the reading-related task--the Oral Reading Task. The Speaking Span Test yielded two types of scores--one strict and one lenient--and, as hypothesized, the two speaking span scores differed in the aspects of fluency they predicted. Speaking span strict correlated better with the Oral Reading Task and the Oral Slip Task, while speaking span lenient correlated more significantly with the Speech Generation Task.

These results are explained by the claim that the Speaking Span Test is a complex measure of working memory span for language production, which taxes both the storage and processing functions of this limited system during the production of speech. While the storage

component of the test is to recall the words presented, the processing component consists of generating grammatical sentences containing these words. Both functions compete for the limited capacity of the system.

Daneman argues that the ability with which an individual coordinates storage and processing in this task is related to his/her ability to produce fluent speech, which also requires efficient coordination of storage and processing of information. It is important to note that the Speaking Span and Reading Span tests are recall tests which were devised to measure working memory span under language production or comprehension processing demands. The tests do not measure processing efficiency per se. Rather, they are assumed to reflect the storage capacity an individual has left as a result of his/her processing efficiency while producing or comprehending language. Thus, as claimed by Daneman and colleagues, good readers have a larger working memory capacity for storing products of the reading comprehension process--such as facts, pronoun referents, and propositions (Turner & Engle, 1989)--because their reading comprehension processing is more efficient and thus they use less of their capacity. By the same token, more fluent speakers have a larger working memory capacity, as measured by the Speaking Span Test, because they are more efficient in executing the processes required during speech production, leaving greater resources available for the storage and subsequent integration of the intermediate products of this processing (Daneman, 1991). The present study examines whether this claim can be made in the case of L2 oral fluency.

II L2 Oral Fluency

We all have an intuitive concept of what it is to be fluent and upon hearing someone talk, we immediately judge him as more (+) or less (-) fluent, although we might not be aware of what makes us consider the speaker as such. Fillmore (1979), one of the first researchers to point out individual differences in fluency, suggests we may judge speakers to be fluent in their L1 in four main ways. In his view, fluency is related to (1) the ability to fill time with fast talk, (2) the ability to produce semantically dense speech, (3) the ability to perform in several pragmatic aspects of language, and (4) the ability to speak with creativity and imagination, building metaphors, punning and making jokes with the meanings and sounds of words, on line.

Highlighting the multidimensional nature of the phenomenon in each of these aspects of fluency, Fillmore proposes that different types of knowledge and skills are involved in the production of fluent speech, mentioning that speakers vary in their vocabulary size, in their knowledge of linguistic forms and formulaic expressions, in their ability to create new

expressions as well as to access and use syntactic constructions of their L1 in the various conversational settings and discourse patterns, varying also in their knowledge of appropriateness of language.

While these four kinds of fluency might well be considered true also for second languages, L2 fluency has been judged and defined in a rather different fashion. As Riggenbach (1991) points out, the notion of fluency has played a much more central role in L2 research than it has in L1, since fluency has been considered an important factor in assessing L2 proficiency.

Most studies dealing with L2 fluency have described the phenomenon at isolated levels of occurrence, from the utterance to the discourse level (Ejzenberg, 1995). As a result, fluency has been defined in a number of different ways. Traditional definitions of fluent speech are "speech that lacks unnatural pauses" or "speech that exhibits smoothness, continuity, and naturalness" (Riggenbach, 1991: 423-24).

In an attempt to organize the ways in which L2 fluency has been understood, Lennon (1990) concludes that the term fluency is generally used in two senses. In its broader sense, it is equated to oral proficiency: a fluent speaker would be the one whose oral production is native-like in all aspects--vocabulary range, grammatical correctness, pronunciation, idiomaticness, appropriateness, and relevance. In its narrower sense, Lennon argues, fluency in an L2 is one component of oral proficiency and is basically related to speech rapidity, to the flow of speech without this being impeded by hesitations. In this narrower sense, fluency is opposed to other components of oral proficiency such as lexical range, grammatical correctness, pronunciation, idiomaticness, appropriateness, and relevance.

This narrower sense is related to the definition Lennon (ibid.) gives for fluency as the perception we have, when hearing someone talk, that the speaker's psycholinguistic processes involved in speech planning and production are working easily and efficiently (p. 391). In line with this view, Schmidt (1992) defines fluency as an *automatic procedural skill* (cf. Carlson, Sullivan, & Schneider, 1989). For him, "fluent speech is automatic, not requiring much attention or effort" (p. 358), in contrast to nonfluent speech, which is effortful and which demands focused attention on a number of processes involved in the various stages of speech production.

Early studies of L2 oral fluency emphasized mainly the temporal variables of speech production. Möhle (1984) compared speech samples of advanced L2 learners of German and French performing a description task and a free discourse task. She was able to identify a number of measures of fluency, among which, speech rate, length and position of unfilled pauses, number and distribution of filled pauses, and length of speech runs between pauses.

Rehbein (1987) analyzed the pauses produced by learners of German as an L2 and developed a set of hypotheses concerning L2 fluency. He posits that fluency is dependent on the activity of planning, which requires the L2 speaker to create a global scheme for his/her utterance. Planning and uttering take place in part simultaneously causing the speaker to pause. Rehbein also points out that fluent speech depends on the type of task the speaker is required to perform, the type of event he/she is involved in, the type of discourse being carried out, and the expectations of the hearer.

Lennon (1990) attempted to quantify the components of fluency by analyzing speech samples of four adult German university students of English as a second language on two occasions--before and after the subjects' study visits to England. Based on the subjects' narration of a sequence of pictures, Lennon devised a wide range of measures of fluency encompassing both temporal variables and disfluency markers, many of them in the tradition of Goldman-Eisler (1968). By comparing each subject's first and last narratives, Lennon found that there had been improvements in their fluency mainly in terms of speech rate and number of filled pauses. He reports that subjects' speech was faster, with fewer repetitions and filled pauses per T-units, less time occupied by unfilled pauses, longer fluent runs between pauses and T-units, and a reduction of pause time at T-unit boundaries.

Riggenbach (1991) is one of the first studies to use conversational data and to include interactive features of speech production in the evaluation of L2 oral fluency. Riggenbach (1991) analyzed the speech of 6 Chinese students of English as an L2, three rated as very fluent, and three as very nonfluent. Her primary goal was to identify which features of the speech of highly fluent nonnative speakers differed from the ones of those considered to be highly nonfluent. Riggenbach asked her subjects to record a dialogue and the quantitative analysis of the speech samples included specific fluency-related items such as hesitation phenomena, repair phenomena, rate and quantity of speech, interactive phenomena, and turn change types. Each of these categories contained a set of sub-items, summing up 19 variables. The results obtained showed few significant differences in features between fluent and nonfluent subjects. However, Riggenbach was able to verify that fluent and nonfluent subjects differed in terms of speech rate and number of filled pauses, supporting Lennon's (1990) findings. Subjects judged as very fluent speakers also showed more ability to make appropriate topic changes and to anticipate the end of turns.

Ejzenberg (1995) investigated the effect of task structure--dialogue vs. monologue--on the display of L2 oral fluency of 50 subjects. In addition, she verified whether there were quantitative and qualitative differences in the speech produced by the very fluent and the very

disfluent subjects. By manipulating the structure of the tasks used in the study, Ejzenberg was able to show that "interactivity" is an important variable affecting speakers' display of fluency (1995, p. 17). Thus, her subjects appeared to be more fluent in dialogues than in monologues, with subjects' fluency varying according to the degree of interactivity present in the context of speech production. The qualitative analysis of four features of speech of three high- and tree low-fluency subjects across tasks showed that high-fluency speakers tend to speak more and faster than their low-fluency counterparts. High-fluency speakers also produce longer talk units and longer fluent units (Ejzenberg, 1995, p. 34 and 36; Postma, Kolk, & Pole, 1990; Pawley & Syder, 1983), displaying, in addition, a number of discourse strategies during speech production in order to maintain an "air of fluency" (Ejzenberg, 1995, p. 38).

Freed (1995) investigated whether native speaker judges' global perceptions of fluency would distinguish between two groups of L2 learners--one with experience in studying in the country of the target language and the other with formal classroom instruction only. Freed also attempted to identify features of fluency that distinguished the two groups. The speech samples of 30 subjects were first subjectively analyzed by a group of 6 native speaker judges on a 7-point scale. Subsequently, linguistic analyses of 8 subjects' speech samples were performed in order to identify attributes of fluency that would help determine those subjects who had been abroad from those who had not. For this linguistic analysis, Freed chose mainly temporal variables and a number of disfluency markers. The analyses performed by native speaker judges' revealed a small difference in the perceived global fluency between the two groups, with a modest increase for the less advanced students (Freed, 1995, p. 134). The linguistic analyses, however, showed that subjects who had lived abroad tended to speak more and faster, with fewer silent and non-lexical pauses, longer speech runs, and a greater number of reformulations and false starts.

The studies reviewed above all focused on L2 fluency as a product of the speech process and attempt to identify the features of L2 fluent speech production. The present study focuses on fluency from a cognitive perspective, thus being primarily concerned with the cognitive processes involved in the production of L2 speech--more precisely, with the ability the speaker has to coordinate the various mental processes involved in this production. For the purposes of the present study, and following Lennon (1990) and Ejzenberg (1995), fluency is here restricted to the oral mode and is considered as a component of language proficiency, being operationalized as the observable speed, accuracy, and fluidity with which speech is delivered (Segalowitz, Segalowitz & Wood, 1998).

III Method

Research hypotheses

The objective of the present study was to verify whether there was a correlation between individuals' working memory capacity and their oral fluency in English as a foreign language (L2)at the discourse level and the articulatory level. A set of experiments was applied in order to assess subjects' working memory capacity and L2 fluency: the Speaking Span Test, in Portuguese and in English, aimed at assessing subjects' working memory capacity; and three tasks aimed at assessing their L2 fluency: a Speech Generation Task, an Oral Reading Task, and an Oral Slip Task. Because oral reading requires, in addition to speech articulation and print decoding, a certain amount of comprehension, the Reading Span Test, in both Portuguese and English, was also included.

Based on Daneman (1991), the present study investigated the following set of hypotheses:

Hypothesis 1: Individuals with a larger working memory capacity as measured by the Speaking Span Test, in Portuguese and in English, would be more fluent at generating speech, more fluent at reading aloud, and less prone to making spoonerisms in the L2.

Hypothesis 2: Speaking Span strict and lenient are sensitive to different aspects of L2 oral fluency: the former would correlate better with fluency in articulation of words, as measured by the Oral Reading Task and the Oral Slip Task, and the latter would correlate better with fluency in producing smooth, continuous, coherent and adequate speech, as assessed by the Speech Generation Task.

Hypothesis 3: The Reading Span Test, like the Speaking Span Test would correlate with fluency in oral reading, but would not correlate with the other two nonreading-related tasks.

Subjects

Subjects for the study were 16 graduate students taking their MA in English language or literature at a major Brazilian university. Of the 16 subjects, 12 were women and 4 were men, ages ranging from 22 to 39 with a mean of 27.5, thus a predominantly young adult sample. At the time of data collection, all of the subjects were working either on their research proposal or on their thesis, and thus had gone through a number of courses which required them to perform in English at a high standard in both the oral and written modes. All of the subjects had previously dealt with English professionally, mainly teaching. Except for three of the subjects, all of them held undergraduate degree in Portuguese/English Languages and Literatures. Subjects' experience in an English-

speaking country varied from short study visits to longer periods of residence, again with the exception of three subjects--not the same three--who had never been abroad. For the purposes of the present study, these subjects are considered to form a relatively homogeneous group in terms of L2 proficiency, sufficient to allow them to use it successfully at least for academic purposes, including speaking. Furthermore, the subjects selected characterize the type of subjects who generally participate in the studies developed in the psychometric correlational approach to working memory, predominantly university students who presumably have more highly developed cognitive skills.

Materials and tasks

Measures of working memory capacity during language production: Subjects working memory capacity for language production was assessed by means of the Speaking Span Test (SST) in Portuguese (SSTP) and in English (SSTE).

Speaking Span Test in English (SSTE): The SSTE was constructed with 40 unrelated one-syllable words, arranged in two sets each of two, three, four, five, and six words. Each word was presented on the middle line of a XT computer video screen for 1 second and was accompanied by a beep. Subjects were instructed to read the words silently. Ten milliseconds after the word had been removed, the next word in the set would appear beside the place the previous word had been presented, on the same line. This procedure was followed, each word slightly to the right, until a blank screen signaled that a set had ended. Subjects were then required to produce orally a sentence for each word in the set, in the order they had appeared and in the exact form they were presented. Thus, for instance, after being presented with the set:

duck pen gas

a subject generated the sentences:

"The duck is in the pond."
"The pen is mine."
"I need some gas."

Subjects were told that there were no restrictions as to the length of the sentences, but they were required to make them grammatical as regards syntax and semantics. After each subject finished generating the sentences for a given set, the next set would be presented and this procedure was followed until all sets had been presented. The two-word sets were presented first, followed by the three-word sets, the four-word sets, and so on. Following Daneman (1991) and Daneman and Green (1986), the measure applied to a subject's speaking span in English in the

present study was his/her total performance in the test, that is, the total number of words for which a grammatical sentence was produced--in this case, the maximum being 40.

Subjects' responses were tape-recorded and, from the analysis of their responses, two types of scores were obtained, as in Daneman (1991): a speaking span strict, when all the grammatical sentences the subject produced contained the target word in its exact form of presentation, and a speaking span lenient, when credit was given for grammatical sentences that contained the target word in a form other than that of presentation (e.g., target word being "dog" and the word in the sentence produced being "dogs"). The main measure of individuals' working memory capacity was the speaking span strict.

There were a few cases in which subjects recalled words out of their order of presentation or in which they inserted or repeated words of previous sets. In these cases, no credit was given for the sentences produced. No subjects produced ungrammatical sentences in terms of syntax and semantics.

The words constituting the SSTE were taken from the word span test used by Harrington and Sawyer (1993) and from the fan test used by Cantor and Engle (1993). The words were randomly organized in the sets, but an effort was made to avoid phonologically similar words in the same set. In order to minimize processing constraints in sentence production and to avoid a possible word-length effect (Baddeley, 1990), this test was constructed only with monosyllabic words, of three to five letters. Despite the fact that this test did not aim at measuring L2 linguistic knowledge, at the end of the test subjects were shown the list of words presented and asked whether there were any words that they did not know or remember the meaning of, in which case the word would be taken out of the subject's responses during the analysis. There were no cases in which a word was unknown to subjects.

Speaking Span Test in Portuguese (SSTP): The SSTP was devised and applied as the SSTE. The only difference between the SSTP and SSTE was that all words in the former were seven letters in length, in replication of Daneman (1991). Daneman's seven-length specification was maintained in order to keep as close as possible to the design of her tests, although number of syllables or spoken duration might have been more adequate criteria, if we assume that what is maintained in memory is acoustic rather than a visual image of the word.

Measures of working memory during language comprehension: Subjects' working memory capacity for language comprehension was assessed by means of the Reading Span Test--RST--(Daneman & Carpenter, 1980, 1983), which was also carried out both in English (RSTE) and in Portuguese (RSTP). According to Daneman (1991), oral reading (one of the tasks aimed at

assessing oral fluency) involves, in addition to fluency of articulation, comprehension processes that are better captured by the RST. Again, as in Daneman (1991), the hypothesis was that the RSTE would correlate to subjects' oral reading fluency in English, as measured by their oral reading time.

Reading Span Test in English (RSTE): The RSTE was constructed with 40 unrelated sentences arranged in two sets each of two, three, four, five, and six sentences. The sentences were adapted from Harrington and Sawyer (1993). Some of them were slightly modified in order to avoid that words contained in the SSTE were repeated as target-words in the RSTE. The sentences were made syntactically simpler and 3 to 4 words shorter than the ones used in Daneman (e.g., 1991), and each one ended in a different word.

Each sentence was presented one at a time on a XT computer screen, and subjects were asked to read them aloud, trying to comprehend them. At the end of a set, when a blank screen appeared, subjects had to recall the last word of each sentence in the set in the order and form they were presented. Instructions were given orally and subjects were explicitly told that this was also a memory test and that they were thus encouraged to recall as many sentence-final words as they could. The time of presentation in this test was not controlled, but depended on the speed of the subjects. The subjects would read each sentence aloud, and as soon as they finished, the experimenter would press the "enter" key on the computer keyboard, causing the next sentence in the set to be presented. This procedure was followed until the 40 sentences had been presented. As in the SSTE, the two-sentence sets were presented first, followed by the three-sentence sets, the four-sentence sets, and so on. Subjects were told that the sets would be increasingly longer.

In order to make sure that subjects were attending to the meaning of sentences and were indeed applying comprehension processes, a grammaticality judgment was included in the test, as in Harrington and Sawyer (1993). This consisted of incorporating one or two ungrammatical sentences in the sets, in initial, middle or final position. The procedures for making sentences ungrammatical in this test were even simpler than the ones suggested by Harrington and Sawyer (ibid.): ungrammatical sentences did not make any sense and had unacceptable subject-verb agreement, unacceptable sequences of nouns and unacceptable verb tenses, thus being ungrammatical both syntactically and semantically.

Subjects were explicitly told that they would find such ungrammatical sentences and were instructed to tell the experimenter when this was the case. They were also told to ignore the ungrammatical sentences during recall of the sentence-final words. For example, in the following set of three sentences, the fourth and ungrammatical sentence was included, which should be

recognized as ungrammatical and its final word not included during recall, which would include, then, only dog, church, and nut[2]:

The young woman and her boyfriend thought they saw a **dog.**

Suddenly the taxi opened its door in front of the **church.**

All that remained in the lunch box was one salted **nut**.

**Car go break stars don't see the house*.*

< blank screen>

In case they did not recognize a sentence as ungrammatical, the experimenter would tell them so before the next sentence was presented.

Again, as in Daneman (1991), a subject's working memory span for language comprehension was his total performance on the test, that is, the maximum number of sentence-final words he/she could recall in the exact order they were presented --in this case 40. There were a few cases in which subjects gave a word out of order word or repeated words from previous sets. In these cases, no credit was given. In the RSTE there was no lenient score, since no subject recalled forms of words other than that in which the word was actually presented.

Reading Span Test in Portuguese (RSTP): The RSTP was applied in the same manner as the RSTE. The test was constructed with 40 unrelated sentences, of 12 to 17 words in length (as in Daneman, 1991), taken from current magazines and newspapers. Each sentence ended with a different word. Similar to the RSTE, to ensure that subjects were attending to sentence meaning, an obstacle to comprehension, rather than a grammatical judgment, was incorporated. This consisted of omitting the Portuguese diacritical marks which distinguish one word from another--the verb *é* (is) from the conjunction *e* (and)--or which facilitate recognition and pronunciation of the word (the cedilla in *açougue*, for instance). Thus, in order to recognize the words and pronounce them correctly, subjects had to make judgments based on comprehension of the rest of the sentence.

Unlike the RSTE, subjects were not told that there would be such sentences in this test and were expected to notice the absence of the diacritics, especially when the meaning was not clear. Instructions were given orally and training was given only to those subjects who found it necessary. Subjects were instructed to recall all sentence-final words of each set presented, in their original form and order of presentation, which would be in increasingly longer sets. The span tests in English--SSTE and RSTE--were included in the present study in order to answer a

---

[2]During the test, to-be-remembered words did not appear in bold and ungrammatical sentences were not preceded by an asterisk.

secondary question: whether working memory capacity, as measured by the speaking span and the reading span tests, remained the same across languages.

Measures of L2 fluency: Subjects' L2 fluency was assessed by means of a Speech Generation Task (SGT), an Oral Reading Task (ORT), and an Oral Slip Task (OST).

Speech Generation Task (SGT): In the SGT subjects were presented with a picture and required to describe it as well as make comments about it for the duration of 1m and 30s. The picture, adapted from an L2 textbook and painted in watercolors on a 20 x 25-cm card, portrayed a detailed scene of a middle-class family at home. In the living room, there were five members of the family, each one doing a different activity. In the kitchen, the family maid was involved with the housework.

Although picture description is generally considered a highly pre-structured task because of the number of cues the speaker has available to organize his/her speech (Ejzenberg, 1995), the scene portrayed in this particular picture left it open to subjects to decide on the gender of one adult character, purposefully not clearly defined in the picture, and on the hierarchical position of a female character, who could either be a mother or a daughter. It was believed that these two aspects of the picture would make the task more demanding.

Subjects were explicitly instructed to give as much information as they could about the picture in their descriptions as well as in their comments. Their speech samples were tape-recorded and then transcribed so that they could be scored. The main measure of fluency was that used by Daneman in her 1991 study, the total number of words produced during the time allotted, or their speech rate.

Subjects' tape-recorded protocols were also submitted to the evaluation of two independent judges, native speakers of American English. These judges, who were also L2 teachers, were asked to rate subjects' fluency in terms of their richness of content, this aspect being subjectively defined by the two judges themselves, on a scale of 1 (repetitious, semantically empty) to 5 (creative, semantically rich).

Since the two independent judges did not receive any training to experimentally assess fluency or detailed instructions about this study, their ratings were given solely on their subjective impression of richness of content as native speakers and as EFL teachers. For this reason, the average rating of the two independent judges was used to correlate with individuals' working memory capacity. It is noteworthy, however, that the main measure of L2 fluency, for the purposes of this study, is the total number of words subjects produced in the time allotted. This second subjective measure provided by the two judges was correlated with subjects' working

memory capacity for L2 fluent speech production in order to answer another secondary question of this research: whether the SST was a good predictor of fluency as directly measured by the listener's subjective impression.

Oral Reading Task (ORT): the ORT consisted of requiring subjects to read aloud a 320 word passage extracted from *The Great Gatsby*, by Scott Fitzgerald. Subjects were told that the emphasis was on reading speed and were explicitly instructed to read the passage as quickly as possible, but not as to slur words. They were also given extra time to read the passage silently first, in order to check for vocabulary and pronunciation problems. Reading time, the measure of fluency, was measured in seconds (s) with a stop watch. Subjects' protocols were tape-recorded.

Oral Slip Task (OST): the OST aimed at eliciting spoonerisms in the laboratory. In her original study Daneman adapted this task from Baars, Motley, and MacKay (1975). The task devised for the present study is different from Daneman's as regards (1) the number of word pairs with which the test was constructed -- 84 in total -- and (2) the word pairs themselves, some of which were collected from Fromkin's (1971) Appendices as well as others devised by the experimenter, since Daneman does not provide Appendices for the items constituting her experiments. In all other aspects, this task was devised as Daneman describes it.

Subjects were presented with the 86 pairs of words on the middle line of a computer video screen, one pair at a time at a rate of 1s each--900 milliseconds (ms) of exposure and 100 ms of interval-- and were required to read them silently (Appendix F). Upon hearing a beep, subjects were to speak aloud the pair which immediately preceded the beep. There were 24 cued (via the beep) word pairs. From these, 8 were target word pairs aimed at eliciting the spoonerism errors, and 16 were filler pairs aimed at disguising the targets. Subjects were expected to speak aloud 24 pairs of words, making 8 spoonerism errors and pronouncing correctly the remaining 16 pairs of words. The measure of fluency in this task was the total number of spoonerism errors the subject made.

Spoonerism errors were induced by provoking a predisposition toward it. This was done by presenting three phonological interference word-pairs immediately before the target word pair. The phonological interference word pairs were similar to the spoonerism expected, therefore biasing the pronunciation of the target word pair. Thus, for instance, subjects were presented with the following pairs of words (one at a time):

sea ships
seek sheeps
see shears
she sees <beep>

On hearing the beep subjects were expected to say "sea shees" instead of "she sees", as the s/sh phoneme pattern in the beginning of words was established by the three phonological interference pairs.

To ensure that subjects would attend to all word pairs, the beep only sounded 500ms after the removal of the to-be-spoken pair from the computer screen, entering 400ms into the 900ms presentation of the next pair. This procedure did not allow subjects to perform the task by ignoring noncued pairs since they did not know when the beep would sound nor which pairs would be beeped.

Subjects were explicitly instructed to attend to each word pair and were also told that the beep would sound after the removal of the to-be-spoken pair. All subjects' protocols were tape-recorded.

Procedures

The data for this study were collected individually with each subject in a room at the university, in two sessions which took place on different days for each subject. There were a computer, a tape recorder and a few chairs in the room. In the first session, subjects' working memory capacity was assessed through the application of the four span tests. In the second session, subjects' L2 fluency was assessed through the other three tasks. Subjects were contacted beforehand to schedule the first session. At the end of the first session of each subject, the second one was scheduled. The interval between one session and the other varied among subjects and depended on the time they had available. Instructions were given orally and in Portuguese--all subjects' L1-- in all of the tests carried out. Subjects were explicitly told that the span tests were memory tests and that it was necessary to focus their attention on the stimuli. Likewise, they were explicitly told what the measures of fluency would be. For the memory tests and the Oral Slip Task, the experimenter first gave sample items. Then, actual previous training was given only if subjects required it and would last until they decided to stop. Instructions were given before the beginning of each experiment and these were applied in the order described above.

Results and discussion

The results obtained from the application of the set of seven experiments were complex and non-systematic across tests. Presentation and discussion of the results will be organized as follows: (1) the span tests; (2) working memory capacity and the Speech Generation Task; (3) working memory capacity and the Oral Reading Task; (4) working memory capacity and the Oral Slip Task.

(1) The span tests: As Table I shows, the Means (M) and Standard Deviations (SD) of the span tests tended to be similar within tests, that is, within the speaking span tests and the reading span tests, irrespective of language.

Table I - Mean Performance and Standard Deviations for Measures of Working Memory Capacity

	Mean	SD
Speaking Span Test in Portuguese (SSTP)-strict	21.5	2.7
Speaking Span Test in Portuguese (SSTP)-lenient	23.3	2.1
Speaking Span Test in English (SSTE)-strict	21.4	2.8
Speaking Span Test in English (SSTE)-lenient	23.4	3.5
Reading Span Test in Portuguese (RSTP)	27	4.7
Reading Span Test in English (RSTE)	24.1	5.1

However, as the results of Pearson Product Moment Correlations show in Table II, only the RS in Portuguese and in English reached significant correlations--[r (16) = 0.78, p = 0.0003]. No significant correlations were found between the SSTP and the SSTE or between the speaking and the reading span tests in either language.

Table II - Correlations among Measures of Working Memory Capacity

	SSTE strict	SSTE lenient	RSTE	RSTP
SSTP strict	0.20	0.16	0.16	0.09
SSTP lenient	0.11	0.27	0.36	0.27
RSTP	0.13	0.17	0.78*	----
RSTE	0.33	0.35	-----	0.78*

* $p < 0.01$

As already observed, the speaking span (Daneman & Green, 1986) and the reading span (Daneman & Carpenter, 1980) tests are complex measures of working memory which require that the individual carry out a processing task while trying to maintain the to-be-remembered stimulus, thus taxing both the processing and storage functions of the system. There are currently three main theories that account for individual differences in working memory capacity: the task-specific view, the processing efficiency view and the activation view (Cantor & Engle, 1993).

The task-specific view posits that the greater an individual's efficiency in processing information, the greater the capacity left available for storage of the products of this processing

and of material retrieved from long-term memory (Cantor &Engle, 1993). This more efficient processing is highly task-specific (Daneman & Green, 1986): an individual's working memory capacity will vary according to his/her efficiency in the processes specific to the task with which working memory capacity is being correlated. Thus, for instance, good readers will have a functionally larger working memory capacity in reading-related tasks, but not necessarily in language production tasks. Within this view, the processing component of the span test must require the same processes present in the cognitive task whose performance is being predicted.

Further elaborations of the task-specific view have led to the processing efficiency view, which claims that there are general skills which are employed in any task demanding the manipulation of language. For instance, Daneman & Tardiff (1987) argue that individual differences in working memory capacity can be measured through processing efficiency alone, without including a simultaneous storage component in the task.

Daneman & Tardiff (1987) examined the relationship between three span tasks (verbal span, math span, and spatial span) and comprehension. The span tasks had both a processing and a storage component. The verbal and math span tasks correlated with verbal abilities. However, to show that the crucial variable in individual differences in working memory is processing efficiency, Daneman and Tardiff added three storage-free span tasks in which only processing was tested. They also found a correlation between these tasks and comprehension which led them to conclude that it is individual differences in processing that explain differences in verbal abilities. Thus, the emphasis is on the efficient processing skills individuals have while performing language-related tasks. The difference between the task-specific and the processing efficiency views is that in the latter the processing component of the span task need not specifically require the same processes of the task being predicted.

The activation view defines working memory as information in long-term memory that is temporarily activated to a level that makes it available for cognitive activity (Cantor & Engle, 1993; Engle, Cantor & Carullo, 1992). The capacity of this system is the total amount of activation an individual has available to retrieve information from long-term memory in order to carry out a cognitive task. Individuals with higher or lower spans, as measured by the span test, differ in their limits of activation. This limited capacity, in the activation view, is independent of the nature of the task, being, thus, a single unitary resource.

It seems reasonable to argue that the results of the span tests carried out in the present study reflect the functional capacity of working memory in relation to the processing requirements made by the background task. Obviously, the background tasks of the speaking and reading span tests involve qualitatively different processes, the former demanding processing

efficiency in language production, and the latter in language comprehension. Furthermore, the lack of a correlation between the SSTP and the SSTE might indicate that speaking in L1 and L2 require somewhat different processing, thus leading to a variation in subjects' working memory capacity in language production, as illustrated by Figures 1 and 2. In fact, the bilingual models of speech production proposed in the L2 literature (e.g., de Bot, 1992; Faerch & Kasper, 1983) differ in some aspects from a unilingual model like Levelt's (1989).

Figure 1-Subjects' performance on the SSTP and SSTE (strict scores)

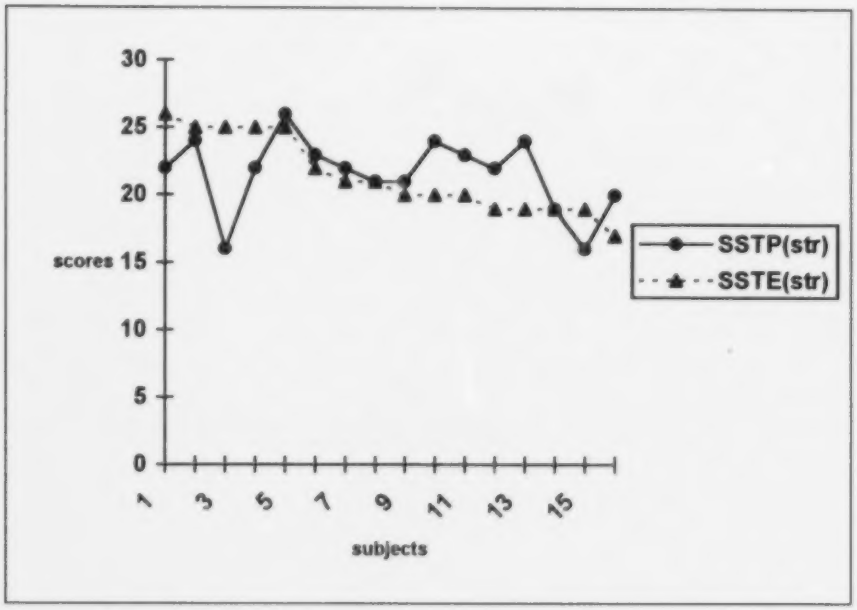
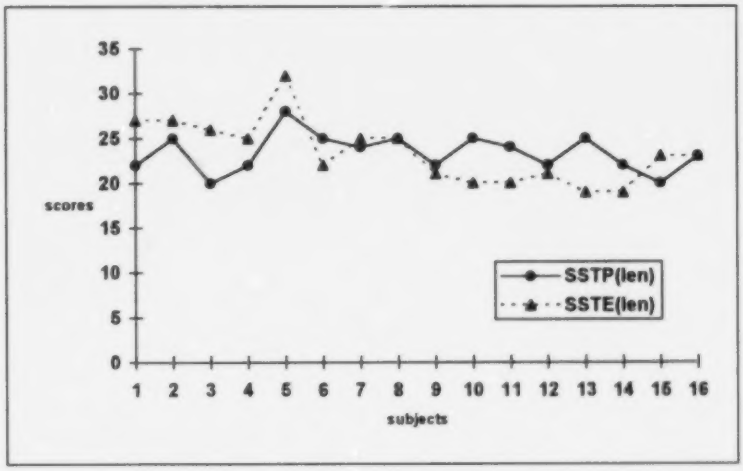


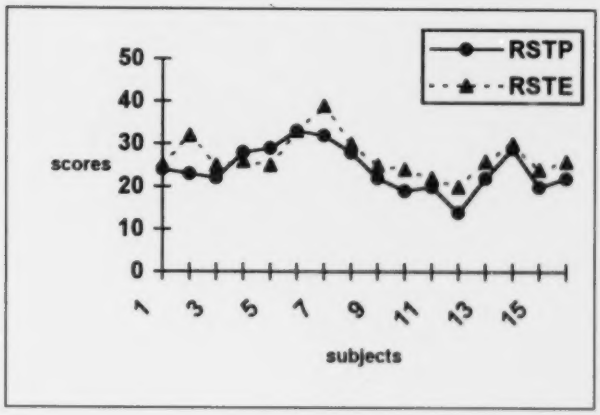
Figure 2-Subjects' performance on the SSTP and SSTE (lenient scores)



The idea that language production in L1 is qualitatively different from that in L2 is also supported by Krashen's (e.g., 1982) well-known acquisition/learning distinction.[3] While speaking in our L1 is a product of the acquisition process, speaking in an L2 is, to a great extent, an outcome of learning, a quite different process. Although the acquisition/learning distinction has been severely criticized, it seems to be in line with the procedural/declarative dichotomy sustained by neuropsychologists (e.g., Paradis, 1994). The type of knowledge that results from the learning processes would be declarative, a conscious, explicit knowledge. It seems reasonable to speculate that, due to the characteristics of L2 learning processes, speaking in the L2 for the subjects who took part in the present study is, at least partially, a learned, rather than acquired, skill. This distinction also seems useful to explain the significant correlation between the RSTP and the RSTE. Reading, unlike speaking, is a product of learning, be it in L1 or L2.

[3] Except where explicitly mentioned that it is Krashen's terminology that is being adopted, as is the case in this section, the terms acquisition/learning have been used interchangeably in the present study.

Figure 3-Subjects' performance on the RSTP and RSTE



Unlike language production, models of reading generally assume that there are no qualitative differences in reading in L1 and L2, which means that, in principle, the processes are the same. The significant correlation found between the RSTP and the RSTE corroborate the similar correlation obtained by Harrington and Sawyer (1993). Figure 3 illustrates, subject by subject, this high positive correlation, and shows that performance on the RSTP was better.

(2) Working memory capacity and the Speech Generation Task: The results of the correlation between working memory capacity, as measured by the Speaking Span Tests (SST), in Portuguese (SSTP) and in English (SSTE), and the Speech Generation Task (SGT), in which fluency was measured in terms of number of words produced, stand as the most important findings of the present study. As hypothesized, the SST correlated significantly with this L2 fluency task, but only the English version of the SST (see Table III). From the scores obtained, the SSTE strict correlated better with L2 fluency in the SGT, although the expectation was that the SSTE lenient would correlate better with fluency in this task [$r(16) = 0.64$, $p = 0.0073$, for the SSTE strict and the SGT; and $r(16) = 0.61$, $p = 0.0114$ for the SSTE lenient and the SGT].

Table III - Correlations among the Speaking Span Tests and the Speech Generation Task

	SSTP strict	SSTP lenient	SSTE strict	SSTE lenient
SGT	-0.08	-0.19	0.64*	0.61**

*$p < 0.01$
**$p < 0.05$

The SGT, following Daneman (1991), was included as a global measure of oral fluency because, to be performed, it requires skillful coordination of the processes involved in the planning and execution of fluent speech. This coordination of the speech production processes is assumed to be carried out in working memory. As Daneman argues, the larger a subject's working memory capacity, as measured by the SST, the more fluent his/her speech will be, since his/her coordination of the speech production processes will also be more efficient. The significant correlation found between the SSTE strict and the SGT corroborates the results of Daneman's study, in which she found a significant correlation between individuals' working memory capacity and L1 oral fluency. However, as already noted, in Daneman's study SST lenient was a more powerful predictor of fluency in the SGT than SST strict.

In all theories of working memory capacity adopted within the psychometric correlational approach, and thus in the present study, working memory is conceptualized as a single central system (Daneman 1991; Daneman & Green, 1986) in charge of the processing and temporary storage of information during complex cognitive tasks. Researchers in the psychometric correlational approach are primarily concerned with what Baddeley (1990) calls the *central executive* and assume no other processing components within this system. Thus, in relating the processes of speech production as Levelt (1989) conceptualizes them to this theory of working memory, it seems reasonable to suggest that the whole process takes place within this single central system, with no particular peripheral components being responsible for specific processes.

Daneman & Carpenter (1980, 1983) proposed that the two functions of working memory, the storage and processing functions, compete for the limited capacity of the system, which acts as an arena for the execution of processes and for the storage of the intermediate products of these processes. Hence, when the individual is engaged in a complex task such as speaking, the capacity of the system is being shared by both conceptual and linguistic processing demands as well as by storage demands. For each of the processes of speech production there is an intermediate outcome that must be temporarily stored.

Working memory is in charge of all the mental processes involved in speaking--from intention to the construction of internal speech. The system resources are shared between the execution of processes--establishment of a communicative intention, conceptualization of the message, formulation of the message through grammatical and phonological encoding--and the storage of the products of these processes--the preverbal message (or conceptual structure), the surface structure (result of grammatical encoding), the phonetic plan of the message (or internal speech). The two speech production processes which are assumed to make the greatest demands on working memory are the establishment of a communicative intention and the

conceptualization of the message (Levelt, 1989), for these two processes, which can hardly be separated, require that the speaker consider aspects of the context in which he/she is involved as being determinant of the kind of talk he/she will produce. Another extremely demanding process is the monitoring of one's own speech, whether as a phonetic plan or as overt speech (Levelt, 1989).

In addition to these higher-level conceptual processes, there are specific linguistic processes, which also compete for working memory processing and storage capacity. These are assumed to be highly automatic in L1, and, for this reason, do not make great demands on the system and thus occur in parallel. Fluent speech requires skillful coordination of all these processing and storage requirements. Daneman (1991) claims that the correlation between subjects' working memory capacity and their fluency in L1 can be accounted for by the fact that individuals with a larger working memory capacity have more efficient processing skills in the task in question--speech production--thus leaving more of their working memory capacity for the storage of intermediate products of this processing and subsequent integration of information processed. Daneman & Carpenter (1983), Daneman & Green (1986) and Daneman (1991) have argued that working memory capacity is task-specific. The results of the present study, with respect to the correlation between the speaking span tests and the Speech Generation Task, tend to corroborate the task-specific view.

In the present study, each subject's working memory capacity measures varied from one span test to another. It seems reasonable to suggest that, at least for speaking, the factor determining different results in two similar tests was the language--speaking different languages imposes different processing and storage demands on the system. The assumption that speaking in different languages involves somewhat different processing and storage requirements is supported by the literature on fluent L2 speech production (see, e.g., de Bot, 1992).

It can be assumed that Levelt's model was proposed primarily for the unilingual speaker since the author makes only modest references to bilinguals. Nevertheless, in view of the tremendous explanatory power of the model, de Bot (1992) has proposed an adaptation of the model for the bilingual speaker, which is useful in explaining how L2 speech production might take place. In the case of the present research, it casts some light on how working memory capacity might be involved in this production.[4]

The first modification de Bot (1992) makes in Levelt's original model in order to adapt it to L2 speech production concerns the *Conceptualizer*, which Levelt assumes is completely

[4] For the purpose of this study, 'bilingual' means the speaker of an L2 independent of the context of acquisition and use of this L2 and level of proficiency.

language-specific. De Bot proposes, instead of considering the Conceptualizer to be completely language-specific, to consider macroplanning--establishing goals and subgoals--to be language-independent, and microplanning--giving structure to the content of goal and subgoals--to be language-specific. This modification seems reasonable, since it is in the microplanning stage that the speaker might be faced with conceptual decisions such as spatial and temporal reference, where the options are different for different languages. Thus, one of the main differences between speaking in L1 or in L2 is that when speaking in the latter, if the individual has organized his thoughts according to the way concepts are expressed in L1, he might have problems in expressing a particular concept for which the L2 does not have the lexical items or whose lexical items the speaker can not access. A native speaker is not normally faced with this difficulty. Problems in the conceptualization of the message in L2 will lead to problems in the formulator when the speaker gets involved in grammatical encoding, whose first step is accessing the specific lexical items to realize the message.

With regard to the formulator, de Bot's proposal is similar to Levelt's in including a separate component for each language. One of the crucial aspects of Levelt's model is the connection between meaning and syntactic information. For Levelt, the speaker first accesses meaning and, based on his/her lexical choices, applies syntactic procedures which are defined by the grammatical specifications of the lexical item in the lemma. Such an assumption poses a relevant question within L2 studies related to the form in which our L2 mental lexicon is organized.

Based on neurolinguistic research, Paradis (1985) claims that bilinguals have one conceptual store and two distinct semantic stores which are differentially connected to the conceptual store. This store, which corresponds to our experiential and conceptual information, contains mental representations of things, events, properties and qualities of objects, and our knowledge of the world. The lexical items representing these concepts are stored separately for each language. Many theoreticians (Grosjean, 1982; Poulisse, 1993; de Bot, Cox, Ralston, Schaufeli, & Weltens, 1995) seem to side with the view that bilinguals have one conceptual store and separate lexical stores for each language.

Having separate lexical stores implies having separate formulators as well. As already noted, it is in the formulator that both grammatical and phonological encoding takes place. De Bot proposes that the bilingual speaker might have separate storages for those language-specific sounds. The articulator, in turn, is assumed to be one and the same for both languages.

The point in giving an outline of how speech production in L1 and L2 occurs is that speaking seems to differ in terms of processing when languages differ. If speaking in different

languages entails different processing, then it seems reasonable to assume that an individual's working memory capacity will vary according to the language being spoken. As Turner & Engle (1989) point out, Daneman and colleagues argue that the working memory capacity span measure--the span test--is dependent on the background task carried out while the span is being measured. This background task must include the processing of the task whose performance the span measure will predict. In other words, if the span measure predicts performance in speaking in English as an L2, the background task (the processing part of the test) must be an activity involving speaking in English. Thus, in principle there seems to be no reason why individuals' working memory capacity as measured by the SSTP (strict or lenient) should be related to their fluency in English, since this test is taxing working memory storage capacity under Portuguese language production processing constraints, a process different from the production of language in English.

As stated above, it was hypothesized, based on Daneman's (1991) L1 findings, that the lenient score of the SST would correlate better with fluency in the SGT, since this task allowed subjects to produce speech in a creative, semantically rich way. This prediction was not born out for the L2. As Table III shows (p.), the magnitude of the correlation for the SSTE strict was higher than that for the SSTE lenient.

Daneman's finding that the lenient score was a better predictor of L1 fluency seems to reflect the assumption that, when presented with the to-be-remembered word, subjects are more inclined to think of the meaning of this word and of the semantic context in which this word may appear in a sentence than to think of the form of this word. Thus, it is possible that, when performing the SST in their native language, subjects are more concerned with making a sentence in which content is the most important aspect and, to conform to this, the word might be slightly modified so that meaning can be better expressed. On the other hand, in the L2, especially if it was learned in the classroom, where they are likely to have been trained in manipulating language, subjects may be inclined to think first of the form in which the word was presented--which restricts its grammatical environment--and then of a possible sentence in which that word could appear. While thinking of the form in which the word was presented, it is possible that subjects become engaged in some type of rehearsal, which would enable them to keep an exact representation of that word in working memory.

In summary, the significant correlations found between the SSTE and the SGT were explained in terms of the task-specific view. In the case of the present study, some individuals have more efficient processing skills for L2 production than others, thus leaving a greater part of their working memory capacity for storage and integration of material. Their L2 speech

production processing efficiency is evident in their performance on the SGT--they were more able to coordinate the planning and execution processes involved in L2 production, which resulted in L2 speech characterized as smooth, continuous, with few perceptible pauses and hesitations, and adequate to the context, as shown by the measure of fluency--number of words produced in the allotted time. Subjects with inefficient processes allocate more of their available capacity to processing, thus leaving less for storage of the to-be-remembered words in their exact form of presentation.

One possible reason subjects with a larger working memory capacity have efficient processing skills in the L2 might be the degree of automaticity of their L2 production processes. In his proposed blueprint for the speech production process, Levelt (1989) assumes that most of the processes involved in speech production take place in parallel incremental fashion, which requires that certain processes be automatic. It is well accepted in the cognitive literature that, because humans are limited-capacity processors, some aspects of the cognitive tasks they are involved in have to be highly automatized so that they can attend to those more complex aspects of the task (McLaughlin, 1987, 1990). The limited capacity of our working memory requires that only some aspects of the task we are engaged in be attended to at a time. These aspects are carried out by means of controlled processes.

The dichotomy between controlled and automatic processes--a classical one within cognitive psychology--is fundamental to the understanding of human cognitive behavior. Shiffrin and Dumais (1981) state that all complex skills involve a mixture of controlled and automatic processes. Automatic processes do not demand processing resources, thus freeing the cognitive system to the more complex, higher-level processing of the task. Because they do not share cognitive resources, automatic processes are highly efficient and can be carried out in parallel. Controlled processes, on the other hand, are assumed to make great demands on the capacity of the system (e.g., Shiffrin & Schneider, 1977).

Levelt proposes that the aspects of speaking that require the most attention are the establishment of a communicative intention and the conceptualization of the message. In other words, when we speak, our attentional resources are focusing on what we want to say. Everytime we speak, we first have to conceive of a message, and it is unlikely that we have a package of intentions stored in our long-term memory. Thus speaking involves an activity of genesis: we create a communicative intention, and, to realize this intention, we have to find an effective means of expression, bearing in mind that we have to be coherent, that we have to develop a chain of thought, that we have to give relevant information, and that the context requires certain social procedures which affect the whole message construction. Another aspect that requires our

attentional resource is monitoring, which happens only if the speaker is aware of what he/she is saying and how he/she is saying it.

Levelt is careful to say, however, that even within conceptualization some aspects are automatized. The adult speaker has such an extensive experience with speaking that many conversational skills are automatized: i.e. packages of ready-made messages or formulaic language are available. The other components of the model proposed by Levelt are assumed to require mostly automatic processing. Formulating the message and articulating it demand very few, if any, controlled processes.

Researchers (e.g., Gatbonton & Segalowitz, 1988; Crookes, 1990; Schmidt, 1992 & 1994; de Bot, 1992; McLaughlin, 1987 & 1990, Paradis, 1985 & 1994, among others) agree that L2 fluency, like fluency in L1, requires a great degree of automaticity, especially in the grammatical and phonological aspects. Such automaticity is necessary in order to free the attentional resources to focus on those more demanding aspects of the task, which Levelt (1989) considers to be the establishment of a communicative intention and the conceptualization of the message. De Bot (1992) argues that it is in the formulator, where the grammatical and phonological encoding takes place, that L2 speakers might face the greatest difficulties, ranging from access of the appropriate lexical items to the application of syntactic and morphophonological rules. A consequence of the lack of automaticity in L2 speech production is slower speech rate and, probably, less creative speech, as they cannot pay as much attention to conceptualization.

It could be argued that it is degree of proficiency, and not necessarily system capacity problems, that is causing individuals' less fluent production. Indeed, one of the most important questions the research within the psychometric correlational approach has not yet answered is whether working memory capacity is dependent on the degree of proficiency the individual has in the task in which this capacity is predicting performance. This problem is even more serious when the complex cognitive task is performed in an L2. Despite all the effort made to diminish differences in subjects' level of proficiency in this study, it is not possible to deny the fact that degree of proficiency is an intervening variable in the results of the present study.

The significant correlation found between working memory capacity as measured by the SSTE and the SGT validates the speaking span test as an instrument to assess working memory during speech production and corroborates the findings of research in L2 fluency at the same time. Lennon (1990), Riggenbach (1991) and Ejzenberg (1995) all reported finding speech rate as measured by words or morphemes produced per minute to be a significant distinctive variable between more and less fluent nonnative speakers. In the present study the main measure of L2

fluency was the number of words produced in the time allotted --1m30s. As observed, more fluent speakers produced more words than less fluent speakers.

In order to check whether working memory capacity would correlate with subjective measures of fluency, subjects' speech samples were submitted to two independent judges to rate their fluency in terms of richness and originality of content on a 5-point scale (1 for repetitious, semantically empty speech and 5 for creative, semantically rich speech). Since these measures are subjective, the average rating of the two judges was used to enter subjects' results. Working memory capacity, as measured by the SSTE, did not reach significant correlations with the SGT as measured by subjective ratings [r (16) = 0.42, p= 0.1018, for the SSTE strict and the subjective measures, and r (16) = 0.35, p = 0.1895, for the SSTE lenient and the subjective measures]. However, there was a significant correlation between number of words produced in the SGT--the main measure of L2 fluency--and the subjective ratings r (16) = 0.54, p = 0.0305]. This result might indicate that number of words produced seems to be taken into consideration when non-trained listeners evaluate L2 fluency on the basis of intuition. Thus, once again, number of words stands as a significant variable in the evaluation of nonnative fluency.

Working memory capacity and the Oral Reading Task: Following Daneman (1991), the oral reading task (ORT) was included in this study primarily in order to assess fluency in articulation of words. The objective of the task was not to evaluate L2 reading comprehension. Nevertheless, it is likely that an individual, when engaged in an oral reading task, is also carrying out reading comprehension processes of the content of the reading passage. For this reason, an index of working memory capacity during language comprehension was also included, that is, the Reading Span Test (RST), which was applied both in Portuguese (RSTP) and in English (RSTE).

It was hypothesized that individuals with a larger working memory capacity would be more fluent in the ORT, thus taking less time to read the whole passage. The prediction was that both the reading and speaking spans, in both languages, would be related to fluency in L2 oral reading, as measured by the time (in milliseconds) each individual took to read the passage, but that the measures of working memory during language production--the SSTP and the SSTE--would be better predictors of fluency in this task, since reading aloud requires the involvement of the speech production processes. It was also expected that the RST would not correlate with the non-reading-related tasks.

The hypothesis was not confirmed. As already noted, the SST in Portuguese did not correlate with any L2 fluency task. The SST in English correlated significantly only with the SGT. Individuals' working memory capacity as measured by the SSTE did not correlate with

fluency as measured by the ORT. Nevertheless, individuals' working memory capacity as measured by the RSTE and the RSTP correlated significantly with the ORT [r (16) = -0.51, p = 0.0455, for the RSTE and ORT, and r (16) = -0.55, p = 0.0263, for the RSTP and the ORT].[5]

There has been a massive amount of research accumulating in the psychometric correlational approach providing evidence that the RST is a good predictor of performance on reading comprehension tasks. However, in such studies the results of subjects' performance on the RST are correlated with measures of specific subskills of reading comprehension, such as making inferences (Mason & Miller, 1983), perceiving inconsistencies in sentences with homonyms (Daneman & Carpenter, 1983), using contextual cues to infer the meaning of new words in a text (Daneman & Green, 1986), to mention only a few.

In the present study, no measures of reading comprehension were applied, but it seems reasonable to suggest that these particular subjects were applying comprehension processes to which the RST was sensitive. If that is true, then it seems also reasonable to suggest that the speed of their oral reading was constrained by reading comprehension processes. In the case of the present results, it seems that subjects were attempting to comprehend in order to read aloud. In fact, it is necessary to have some comprehension of the passage in order to make appropriate pauses and use appropriate intonation while reading aloud. The only reason I see for these subjects to have had their reading time constrained by their reading comprehension processes to such a degree that the SSTE was not sensitive to their L2 oral fluency is the fact that the reading passage chosen for the experiment was not an appropriate one. The passage is decontextualized in the sense that it was taken from the middle of a chapter, it does not have the pattern of a common reading passage, i.e., beginning, middle and end, and the vocabulary is characteristic of literary prose. The strangeness of the passage might have activated higher-level reading comprehension processes in subjects which could be captured by the RST.

An interesting aspect of the results of the correlation between working memory capacity and the ORT was that the RST both in Portuguese and English correlated significantly with this task. These results are in line with previous findings of significant correlations between working memory capacity for reading in L1 and for reading in L2 (Harrington & Sawyer, 1993; Osaka & Osaka, 1992; Osaka, Osaka, & Groner, 1993). As is well known, models of reading generally assume that there are no qualitative differences in reading in L1 and L2. That is not true for language production, since researchers have argued for bilingual models of speech production, emphasizing that the processes are qualitatively different.

---

[5]Shorter reading times in the ORT mean more fluent oral reading, thus the negative correlations.

Working Memory Capacity and The Oral Slip Task: Finally, the last set of experiments in the present study consisted of correlating working memory capacity with L2 speech errors, assessing fluency at the articulatory level. Following Daneman (1991), the Oral Slip Task was used in this study to elicit spoonerisms in an artificial context. A spoonerism is a type of speech error which consists of the exchange of phonemes in adjacent or near-adjacent syllables or words (Motley, Baars, and Camden, 1983). The OST assesses fluency in the articulation of individual words. The hypothesis was that individuals with a larger working memory capacity, as measured by the SST, would be less prone to making spoonerisms in the L2.

No significant correlations were found between individuals' working memory capacity and L2 spoonerisms in either form of the SST. These results will be explained in terms of a methodological failure.

The OST utilized by Daneman (1991) was an adaptation of a technique provided by Baars et al. (1975) to elicit speech errors in the laboratory, the SLIP (Spoonerisms of Laboratory Induced Predisposition) paradigm. Baars and colleagues have extensively used the SLIP technique to evaluate frequency and type of spoonerisms, and the consistency of their findings has led them to classify the SLIP technique as "robust" (Motley, Baars, & Camdem, 1983). However, Sinsabaugh & Fox (1986) point out that in reviewing relevant literature on spoonerisms, they found no published replication of the SLIP paradigm, with the exception of the one produced by Baars and colleagues themselves. Sinsabaugh and Fox (1986) utilized the Baars et al. paradigm in an attempt to elicit spoonerisms in the laboratory and found that the kinds of error that occur are a result of memory confusions rather than the elicitation of real spoonerisms.

The design of the OST task used in the present study is basically the same as that used by Daneman (1991) and by Sinsabaugh and Fox (1986), which in turn are similar to the original SLIP paradigm. The differences between the OST of the present study and the others is that the word pairs of the test are in the L2 and the total number of word-pairs included was smaller. The results of the present study are in line with the L1 results obtained by Sinsabaugh and Fox (1986), who report that spoonerisms were only a small fraction of the errors their subjects made. Much more common were errors involving no spoken responses, responses that were phonetically unrelated to any beeped word pairs, and responses that were phonetically unrelated to the target-word pairs, among others. In the present study the most common type of error made was no spoken response.

Interestingly, when all types of errors, including spoonerisms, other speech errors, and no spoken responses, are indistinctly entered for statistical computation, a significant correlation

between the SSTE strict and the OST is obtained [r (16) = -0.58, p= 0.0182]. That is, individuals with a larger working memory capacity, as measured by the SSTE strict perform better on the OST. It is not possible to say from these results, though, that these individuals are less prone to making spoonerisms. Indeed, individuals with larger working memory spans tended to respond to all beeped word pairs, without a miss, in addition to producing the word pairs correctly. Individuals with smaller spans, although often making no spoonerism errors either, tended to give no spoken responses for some of the beeped pairs or give other words as responses. Thus, it seems that working memory capacity is playing a role in the performance of the task--which may somehow measure English articulation ability--although it is not possible to comment on the role of this system in the production of spoonerisms, due to the methodological problem described.

IV Limitations of the Study and Suggestions for Further Research

Given its complex multidimensional nature, L2 oral fluency has been approached from various perspectives, which resulted in a fragmented view of what it means to be fluent in an L2. Most studies have assessed L2 fluency at a single level of occurrence (Ejzenberg, 1995) and have concentrated on the comparison of high-fluency and low-fluency speakers or on the comparison of native and nonnative fluency. These studies have, in general, examined temporal and linguistic aspects of L2 speech production, with a few attempts to add to these two categories the functional importance of pragmatic features (Riggenbach, 1990), the effect of context on the display of fluency (Ejzenberg, 1995) and the effect of the study abroad context on oral fluency development (Freed, 1995). The present study examined fluency from the perspective of information processing theory by claiming that individual differences in working memory capacity might be related to the production of fluent L2 speech at the discourse and articulatory levels. However, being tentative and exploratory in nature, a number of difficulties were encountered throughout the research in both the theoretical and practical aspects.

First, the assessment of fluency was limited to one speech generation task--in which only one quantitative measure was used--and two other tasks eliciting fluency at the articulatory level. In all three tasks, the point in focus was speed and accuracy of speech production. Further research might investigate the correlation between working memory capacity and L2 fluent speech production as measured by temporal and linguistic variables, in a qualitative, rather than purely quantitative, approach. Second, the sample size used in the present study does not make it possible to generalize the results obtained to other populations of L2 learners. Further research might investigate the relationship between working memory capacity and L2 speech production in larger, more representative samples of L2 learners. Third, although the results of the present

study seem to support the task-specific view of working memory capacity, it is important to point out that the nature of individual differences in working memory capacity is still an unresolved issue. A greater effort is necessary to determine whether this is a functional task-specific capacity or a general capacity underlying performance across tasks. Future research might address this issue and, through the use of various methods to measure working memory capacity and the assessment of various L2 skills, shed some light on processing efficiency, storage capacity, and variation in performance.

References

ASHCRAFT, M.H. (1994). *Human memory and cognition*. New York: Harper Collins.

BAARS, B.J., MOTLEY, M.T., & MAcKAY, D.G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 14*, 382-391.

BADDELEY, A.D. (1990). *Human memory: Theory and Practice*. Hove, UK: Lawrence Erlbaum

BADDELEY, A.D. (1992c). Working memory: The interface between memory and cognition. *Journal of Cognitive Neuroscience, 4*, 281-288.

BADDELEY, A.D. (1999). Essentials of human memory. Hove: Psychology Press.

BADDELEY, A.D. & HITCH, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-90). New York: Academic Press.

CANTOR, J., & ENGLE, R. (1993). Working memory capacity as long-term activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1101-1114.

CARLSON, R.A., SULLIVAN, M., & SCHNEIDER, W. (1990). Practice effects and composition: A reply to Anderson. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 531-533.

CROOKES, G. (1991). Second language speech production research: A methodologically-oriented review. *Studies in Second Language Acquisition, 11*, 367-383.

DANEMAN, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research, 20*, 445-464.

DANEMAN, M., & CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal or verbal learning and verbal behavior, 19*, 450-466.

DANEMAN, M., & CARPENTER, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 561-584.

DANEMAN, M., & GREEN, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25*, 1-18.

DANEMAN, M. & TARDIFF, T. (1987). Working Memory and reading skill re examined. In M. Coltheart (Ed.), *Attention & performance* (Vol. 12, pp. 491-508). Hillsdale, NJ: Lawrence Erlbaum.

De BOT, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics, 13*, 1-24.

De BOT, K., COX, A., RALSTON, S., SCHAUFELI, A., WELTENS, B. (1995). Lexical processing in bilinguals. *Second Language Research*, 11, 1-19.

DECHERT, H.W. (1983). How a story is done in a second language. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 175-195). London: Longman.

DECHERT, H.W. & RAUPACH, M. (Eds.). (1980).*Temporal variables in speech*. The Hague: Mouton.

DECHERT, H.W. & RAUPACH, M. (1987). *Psycholinguistic models of production*. Norwood, NJ: Ablex.

EJZENBERG, R. (1995). *Discourse markers and context-dependency in oral fluency assessment*. Paper presented at the 1995 American Association of Applied Linguistics Conference, Long Beach, California.

ENGLE, R. (1996). Working memory and retrieval: An inhibition-resource approach. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus, R.T. Zacks, (Eds.), *Working memory and human cognition* (pp. 89-119). New York: Oxford University Press.

ENGLE, R.W., CANTOR, J., CARULLO, J.J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972-992.

FAERCH, C., & KASPER, G. (1983). Plans and strategies in foreign language communication. In C. Faerch and G. Kasper (Eds.), *Strategies in interlanguage communication* (pp.20-60). London: Longman.

FILLMORE, C.J. (1979). On fluency. In C. J. Fillmore, D. Kempler, & W. S. Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp.85-102) New York: Academic Press.

FITZGERALD, F.S. (1926). *The Great Gatsby*. Middlesex, England: Penguin Modern Books.

FREED, B. (1995). What makes us think that students who study abroad become fluent? in B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Amsterdam: John Benjamins.

FROMKIN, V. (1973). *Speech errors and linguistic evidence*. The Hague: Mouton Publishers.

GATBONTON, E., & SEGALOWITZ, N. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly, 22*, 473-492.

GOLDMAN-EISLER, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.

BEST COPY AVAILABLE

GROSJEAN, F. (1982). *Life with two languages*. Cambridge, MA: Harvard University Press.
HARRINGTON, M., & SAWYER, M. (1993). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*, 25-38.
HITCH, G., & BADDELEY, A.D (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology, 28*, 603-21.
LENNON, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*, 387-417.
KRASHEN, S. (1982). Principles and practice in second language acquisition. New York: Pergamon.
LEVELT, W.J.M. (1989). *Speaking. From intention to articulation*. Cambridge, MA: Bradford/MIT Press.
LOGIE, R.H. (1996). The seven ages of working memory. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus, R.T. Zacks, (Eds.), *Working memory and human cognition* (pp. 31-65). New York: Oxford University Press.
MASSON, M.E.J., & MILLER, J.A. (1983). Working memory and individual differences in comprehension and memory of text. *Journal of Educational Psychology, 75*, 314-318
McLAUGHLIN, B. (1987). *Theories of second language learning*. London: Edward Arnold.
McLAUGHLIN, B. (1990). Restructuring. *Applied Linguistics, 11*, 113-128.
MIYAKE, A. & FRIEDMAN, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp.339-363). Hillsdale: Lawrence Erlbaum.
MIYAKE, A., JUST, M.A., & CARPENTER, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language, 33*, 175-202.
MÖHLE, D. (1984). A comparison of the second language speech production of different native speakers. In H. -W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp.26-49). Tübingen: Narr.
MOTLEY, M.T., & BAARS, B.J. (1976). Laboratory induction of verbal slips: A new method for psycholinguistic research. *Communication Quarterly, 24*, 28-34.
MOTLEY, M.T., BAARS, B.J., & CAMDEN, C.T. (1983). Experimental verbal slip studies: A review and an editing model of language encoding. *Communication Monographs, 50*, 79-101.
OSAKA, M., & OSAKA, N. (1992). Language-independent working memory as measured by Japanese and English span reading span tests. *Bulletin of the Psychonomic Society, 30*, 287-289.
OSAKA, M., OSAKA, N, & GRONER, R. (1993). Language-independent working memory: Evidence from German and French reading span tests. *Bulletin of the Psychonomic Society, 31*, 117-118.
PARADIS, M. (1985). On the representation of two languages in one brain. *Language Science, 7*, 1-39.
PARADIS, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism. In N. Ellis (Ed.), *Implicit and explicit language learning* (pp. 393-419). New York: Academic Press.
PAWLEY, A., & SYDER, F.H. (1983). Two puzzles for linguistic theory: Nativelike selection and native fluency. In J.C. Richards & R.W. Schmidt (Eds.), *Language and communication* (pp. 191-226). New York: Longman.
POSTMA, A., KIOK, H. & POVEL, D.J. (1990). On the relation among speech errors, disfluencies, and self-repairs. *Language and Speech, 33*, 19-29.
POULISSE, N. (1993). A theoretical account of lexical communication strategies. In Schreuder, T., & Weltens, B. (Eds.), *The bilingual lexicon* (pp. 157-189). Amsterdam: John Benjamins.
REHBEIN, J. (1987). On fluency in second language speech. In H. -W. Dechert & M Raupach (Eds.), *Psycholinguistic models of production* (pp. 97-105). Tübingen: Narr.
RICHARDSON, J.T.E. (1996). Evolving concepts of working memory. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus, R.T. Zacks, (Eds.), *Working memory and human cognition* (pp. 3-30). New York: Oxford University Press.
RIGGENBACH, H. (1991). Toward an understanding of fluency: a microanalysis of nonnative speaker conversations. *Discourse Processes, 14*, 423-41.
SCHMIDT, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.
SCHMIDT, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 423-41.

SEGALOWITZ, S.J., SEGALOWITZ, N., & WOOD, A. (1998). Assessing the development of automat ity in second language recognition. *Applied Psycholinguistics, 19*, 53-67.

SHIFFRIN, R.M., & DUMAIS, S. (1981). The development of automatism. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp.111-140). Hillsdale: Lawrence Erlbaum.

SHIFFRIN, R.M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing II: Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127-190.

SINSABAUGH, B. A., & FOX,R.A. Reevaluating the slip paradigm: A research note. *Communication Monographs, 53*, 335-341.

STILINGS, N., FEINSTEIN, M., GARFIELD, J., RISSLAND, E., ROSENBAUM, D., WEISLER, S., & BAKER-WARD, L. (1987). *Cognitive science: An introduction*. Cambridge, MA: the MIT Press.

TOMASELLO, M., & HERRON, C. (1991). Experiments in the real world: A reply to Beck and Eubank. *Studies in Second Language Acquisition, 13*, 513-517.

TOMITCH, L.M.B. (1995). *Reading: Text organization perception and working memory capacity*. Unpublished doctoral thesis, Universidade Federal de Santa Catarina, Florianopolis.

TURNER, M. & ENGLE, R. (1989). Is working memory capacity task-dependent? *Journal of Memory and Language, 28*, 127-154.